

Verschmelzungsformen in German: A Lexical Analysis

RENATE RAFFELSIEFEN

University of Washington

1. THE DATA

In Standard German there are alternations between reduced and full forms of the definite article in examples such as the following:

- | | | |
|-----|-----------------------------|--------------------------------|
| (1) | Max geht zur Schule. | Max geht zu der Schule. |
| | Max schreibt ans Amt. | Max schreibt an das Amt. |
| | Max war beim Friseur. | Max war bei dem Friseur. |
| | Max stellt es untern Tisch. | Max stellt es unter den Tisch. |

Schaub regards the forms in the lefthand column as "satzphonetische Erscheinungen" [12:89] that can be explained on articulatory grounds. However, phonological factors are only relevant with regard to the inventory of standardized contracted forms (cf. 1.1.1), but have no bearing on their status in the grammar. The contracted forms are neither due to fast nor casual speech since there is no phonological process in German deleting syllable-initial *d* or *dV*, nor are they in free variation with the full forms as we find a divergence in distribution and meaning. Before we address the question of how to analyze the contracted forms, we will investigate phonological, morphological as well as semantic conditions on their admissibility in Standard German.

1.1. THE INVENTORY OF FORMS

1.1.1. Gaps Due to Phonological Constraints

Although a purely phonological account for the contracted forms in (1) is ruled out, phonological factors should be taken into consideration in order to explain certain gaps in the inventory of Verschmelzungsformen in standard speech. Whereas *beim* occurs even in formal or elevated speech, the forms listed in (2) appear to be somewhat casual and are avoided in formal speech styles.

- | | | | |
|-----|---------|---------|-------------|
| (2) | mit'm | [mItm] | 'mit dem' |
| | durch'n | [dUrçn] | 'durch den' |
| | aus'm | [awsɐm] | 'aus dem' |
| | als'n | [alsn] | 'als den' |

The marginal acceptance of these forms in standard speech is related to the fact that they violate phonotactic constraints. One can certainly distinguish between different degrees of acceptability depending on how severe the violation of a given surface phonetic constraint is, cf. ?*aus'm* (*aus dem*) vs. **aus'r* (*aus der*). In what follows I will not have much more to say about this kind of gap in the inventory of Verschmelzungsformen.

1.1.2. Gaps Due to Morphological Constraints?

Although the nonoccurrence of certain forms and the relegation to substandard speech of others is due to pronounceability, this is certainly not always the case. One cannot possibly claim that the contracted forms in (3) are ruled out on phonological grounds.

- (3) Er war immer gut **zun* Kindern. (zu den) cf. *zum*
 **Bein* Olympischen Spielen treffen sich Sportler aus aller Welt.
 (bei den) cf. *beim*

The only generalization that holds for these cases is that the articles in question are always plural. They could not be accusative singular since *bei* and *zu* are "dative-prepositions."

A second restriction on the admissibility of Verschmelzungsformen involves the genitive singular masculine/neuter determiner *des*.

- (4) **statt's* Filmes (statt des)
 **wegen's* Geldes (wegen des)
 **innerhalb's* Kreises (innerhalb des)
 **während's* Tages (während des)
 **aufgrund's* Regens (aufgrund des)

Again, no phonotactic constraints are violated, and yet these forms are ungrammatical.

Finally, the article *die* (fem., sg., nom/acc.) never appears in Verschmelzungsformen. While it is true that only final consonants of definite articles appear in combination with prepositions, this does not really explain why forms like **aufie* (*auf die*), **fürie* (*für die*), etc., are not possible.

Although one can clearly state in morphological terms the restrictions mentioned thus far, it is not clear what it would mean to claim that they are morphologically "conditioned." Thus, although we can explain the nonoccurrence of **ausr* for example, we cannot yet explain forms like **bein*. Noting that plural is marked, as opposed to singular, and therefore is not subject to contraction, does not strike one as an enlightening explanation. I will return to these "mysterious" gaps in (2.3.2).

1.2. SYNTACTICALLY AND/OR SEMANTICALLY CONDITIONED RESTRICTIONS

Probably the most important argument against the view that Verschmelzungsformen are due to phonologically conditioned reduction-processes is related to the observation that in certain constructions only the contracted forms are grammatical. This issue has been addressed by a number of grammarians;¹ however, the discussions of the restrictions to be found usually amount to a more or less random enumeration of the respective contexts. Thus it has been noted that in many idiomatic expressions only the contracted forms occur.

- (5) Man sollte diese Gelegenheit beim Schopfe fassen. (*bei dem Schopfe)
 Er ist am Ende. (*an dem Ende)
 Man sollte diese Möglichkeit ins Auge fassen. (*in das Auge)
 Man hat ihn zum Narren gehalten. (*zu dem Narren)
 Er kam vom Hölzchen aufs Stöckchen. (*von dem Hölzchen, *auf das Stöckchen)
 Er ist das fünfte Rad am Wagen. (*an dem Wagen)
 Man hat ihn zur Rede gestellt. (*zu der Rede)
 Er sollte sich nicht aufs hohe Roß setzen. (*auf das hohe Roß)

The same is true for nominalized infinitives:

- (6) Beim Essen spricht man nicht. (*bei dem Essen)
 Er geriet ins Stocken. (*in das Stocken)
 Ich finde das zum Lachen. (*zu dem Lachen)
 Sind die Grünen wieder im Kommen? (*in dem Kommen)
 Ich halte nichts vom Fernsehen. (*von dem Fernsehen)
 Er war gerade im Schreiben begriffen. (*in dem Schreiben)
 Das hindert ihn am Arbeiten. (*an dem Arbeiten)

¹ cf. [4:491ff]; [9:134ff]; [11:36ff].

Likewise for nominalized adjectives:

- (7) Im Geheimen wundere ich mich über ihn. (*in dem Geheimen)
 Im Dunkeln fürchtet er sich. (*in dem Dunkeln)
 Er verliebt sich immer wieder aufs Neue. (*auf das Neue)

In fact, in absence of any further syntactic restriction, only the contracted forms are idiomatic:

- (8) Ich war beim Arzt. ?Ich war bei dem Arzt.
 Ich war im Kino. ?Ich war in dem Kino.
 Ich gehe zur Schule. ?Ich gehe zu der Schule.

Ich war bei dem Arzt or *Ich war in dem Kino* elicit the questions *Bei welchem Arzt?* and *In welchem Kino?* As might be expected, these sentences are fine with a restrictive relative clause:

- (9) Ich war bei dem Arzt, den du mir empfohlen hast. (. . . ?beim Arzt, den . . .)
 Ich war in dem Kino, das nur russische Filme zeigt. (. . . ?im Kino, das . . .)
 Ich gehe zu der Schule, die am nächsten liegt. (. . . ?zur Schule, die . . .)

Another way of licensing the noncontracted forms is by modifying the noun by either an adjective or a prepositional phrase:

- (10) Ich war bei dem neuen Arzt. (. . . ?beim neuen Arzt)
 (11) Ich war in dem Kino neben dem Bahnhof. (. . . ?im Kino neben dem Bahnhof)

The contracted forms seem to be somewhat ungrammatical in contexts (9)-(11); Duden comments as follows:

“Der bestimmte Artikel steht vor allem dann getrennt, wenn das Folgende durch einen Gliedsatz oder durch den Rede- oder Textzusammenhang näher bestimmt wird. . . . Der Artikel wird dann selbständig gebraucht, wenn nicht allgemein ein Bereich, sondern etwas einzelnes, näher Bestimmtes, Bekanntes angesprochen wird [4:493]

The restrictions observed here account for the grammaticality judgments with regard to the data listed in (5)-(7) as well. Thus, idiomatic expressions

do not allow further modification through relative clauses, adjectives or prepositional phrases, and the same is true for the nominalized adjectives and the infinitive constructions where "Vorgang oder Zustand" appears to have no "zeitliche Begrenzung" [4:492]. The inherently unrestricted nature of these constructions is apparently not compatible with any kind of modification.

Hinrichs [7:129] talks about the generic vs. nongeneric use of the definite article in this context, thereby nicely accounting for the data given above:

"... the internal structure of the definite generic is limited in large numbers of ways (all relative clauses, most adjectives, and all AP-phrases and all PPs are disallowed). . . ." ² [3:280]

It is not clear if this term is really appropriate since by saying *Ich war beim Arzt* it is, of course, implied that one talks about a particular physician which might even be clearly identifiable in a given context. On the other hand, this particular physician is only referred to in his capacity as a member of a certain class. By way of contrast, consider the following examples:

- (12) Die Amerikaner flogen als erste zum Mond. (*zu dem Mond)
 Köln liegt am Rhein. (*an dem Rhein)
 Er ist am vierzehnten Oktober 1928 geboren. (*an dem vierzehnten Oktober)
 Er hat nur dummes Zeug im Kopf. (*in dem Kopf)

According to Duden, we should clearly expect the noncontracted forms here since it is always the case that "... etwas einzelnes, näher Bestimmtes, Bekanntes angesprochen wird . . ." (see above). However, familiarity with and uniqueness of the objects referred to seem to be prerequisites for the use of the contracted forms here. In this connection, the discussion in Sanders is relevant:

"Gewöhnlich auch: Die Sterne am Himmel. An dem Himmel würde diesen Himmel gleichsam von einem andern unterscheiden, z. B: Der

² I am not convinced that appositive relative clauses are incompatible with generic determiners. Rather, only restrictive relative clauses are unacceptable in combination with Verschmelzungsformen:

"Dagegen ist wieder die Aufstellung der Sprachlehrer, daß ein sich an ein Hauptwort anschließender Relativsatz vor jenem den vollen Artikel nötig mache, in dieser Allgemeinheit nicht treffend, sondern nur für diejenigen Relativsätze, die das Hauptwort nach einer besonderen Art hin bestimmen und erläutern, nur also wenn der Artikel vor diesem stärkere hinweisende Kraft hat." [9:135]

- (12) ... schönste Stern an dem Himmel meines Glücks, ... Die Dorfbewohner sprechen von ihrer Weide am Dorf, die Regierung auch von der Weide an dem Dorf(e), ...” [11:37]

The question then arises of whether we should list the contexts in which the contracted and noncontracted forms occur, or if we should look for a property which would account for the observed complementary distribution. It is interesting to note that the German dialect of Amern displays two distinct sets of definite articles, the distribution of which exactly parallels the distribution of the contracted vs. noncontracted forms.³ This would be rather odd if the respective distribution was contingent upon an enumeration of unrelated contexts. The essential observation is that Verschmelzungsformen are incompatible with any type of restrictive modification such as an adjective, relative clause, or prepositional phrase. This has already been mentioned with regard to the inherently unrestricted nouns given above, but the same is true for the cases listed in (12), albeit for different reasons. Thus, the reader may convince himself that it is odd for pragmatic reasons to further specify these nouns. This is because they are familiar and unique, hence specified to a degree that any further specification would be out of place.⁴ I am not aware of any grammatical term that exactly covers the distributional patterns observed here. For the time being, I will adopt the term ‘generic,’ keeping in mind that I use it for something like ‘nonmodifiable,’ ‘nonspecifiable,’ ‘nonrestrictable.’

Finally, two additional restrictions should be pointed out: First, if the definite article is used anaphorically, contraction is not allowed:

- (13) In der Stadt gibt es ein Kino_i und ein Theater.
 Gestern war ich in dem Kino_i.
 *Gestern war ich im Kino_i.

Second, if the definite article is used deictically, the contracted forms are also not allowed. In this case, the article is always stressed:

- (14) Wie kann man nur bei dém Wetter ausgehen!
 *Wie kann man nur béim Wetter ausgehen.

³ There is a detailed description of the distribution of these articles in Heinrichs [7:85ff].

⁴ It does not always seem to be true that attributive adjectives are incompatible with Verschmelzungsformen. Thus, we find expressions such as:

- aufs hohe Roß (cf. 7)
- am Heiligen Abend

Note, however, that in these cases the adjective and the following noun form a monosemic expression; the adjective, thus, does not have any restrictive function.

Summarizing, we can say that the contracted forms may be used only 'generically,' in which case their use is obligatory.

One further case of allegedly syntactic restrictions on the use of contracted forms is discussed in Hinrichs [7:130]. He argues that the forms *am* and *im* in superlative constructions of adjectives and adverbs are obligatorily contracted from *an dem* and *in dem* respectively. This would account for the following grammaticality judgments:

Gottlieb schwimmt am schnellsten.

*Gottlieb schwimmt an dem schnellsten.

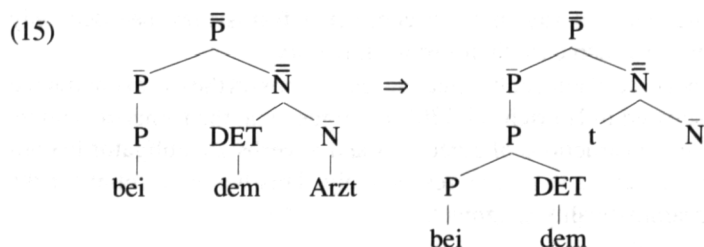
However, it is by no means clear than *im* and *am* can be related to *in dem* or *an dem*. In fact, both the prepositions *in* and *an* and the sg/dat/masc/neut definite article are completely out of place in superlative constructions. The only conceivable argument in favor of this analysis is the reference to the homophonous forms *im* and *am* as used in (1) which, of course, does not mean anything. In my opinion, the only reasonable analysis is to regard *im* and *am* as separate lexical entries which select superlatives.

2. ANALYSES

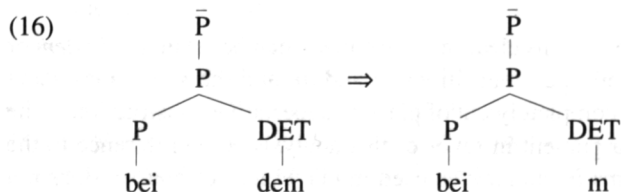
In the following section I will discuss different analyses for the Verschmelzungsformen in German that have been proposed in the past few years.

2.1 CLITICIZATION

Zwicky's 1977 paper [14], intended as a pretheoretical study, has given rise to the current assiduously debated theory of cliticization. Zwicky proposed drawing a distinction between simple and special clitics, the main difference between these two categories being that the latter show special syntax, whereas the former occur in the same linear positions as the full forms [14:6]. According to this taxonomy, the Verschmelzungsformen in German appear as potential candidates for simple cliticization, which indeed has been proposed as an analysis by Zwicky [14], [15]. The formation of the contracted forms comes about in two stages according to this theory: First, a restructuring rule Chomsky-adjoins the 'future-clitic' to its host, and then a rule of allomorphy inserts the reduced variant. In order to generate a form like *beim*, the article must first be adjoined to the preceding preposition.



The reduced variant of the allomorph inserted will always be the final consonant of the 'original' article; in this case it would be an *m*.



The second stage seems to be the appropriate level in the grammar where certain forms can be ruled out, (cf 1.1.1. and 1.1.2). But whereas, depending on where syllabification applies, the phonologically conditioned gaps presumably don't cause too much trouble, the morphologically 'conditioned' gaps most definitely do pose a problem: the nonoccurring forms would have to be ruled out using idiosyncratic features on the articles *des*, *die* and all the plural articles. The semantic differences between the contracted and the non-contracted forms—the generic vs. the anaphoric or deictic use of the definite article—could be accounted for considering the fact that the reduced forms are not derived from the full forms by phonological rules, but are allomorphs which are listed in the lexicon. Divergence in meaning and distribution is therefore expected (cf. Kaisse [8:42]). With regard to the occurrence of *is* vs. *'s* in colloquial English, for example, it has been noted that *'s* may occur in certain constructions with plurals, whereas the unreduced form *is* may not. The clitic has thus apparently "a different range of occurrence than the full form, not merely a more circumscribed one" Kaisse [8:43].

However, the main problem which remains to be solved, is to state the conditions under which cliticization takes place. Kaisse [8] investigates several cases of simple cliticization in English in the Government and Binding framework and states restrictions by referring to constituent structure:

"The fact that so many simple cliticizations make use of c-command

and its stronger relative, government, cannot be coincidental. On the basis of the five English examples treated here, I make the hopefully not premature claim that all simple cliticizations in all languages are restricted to applying between terms bearing some c-command relation to one another." [8:73]

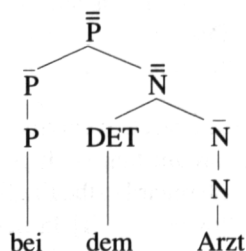
The notions of government and c-command that she adopts in her analysis are defined as follows:

"The head of any phrase will be said to govern all the phrases ($=X^{\max}$) within its projections, and to c-command every element within those phrases." [8:47]

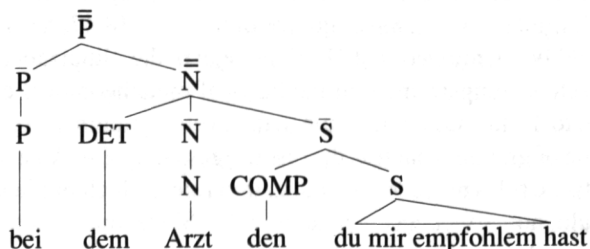
As I will show in a moment, it is not possible to account for the distribution of the Verschmelzungsformen as opposed to the noncontracted forms in German in this fashion. This suggests that either Kaisse's generalization quoted above is indeed premature, or that the contracted forms in German are not instances of simple cliticization.

To begin with, it should be noted that potential syntactic restrictions should be found within the $\bar{\bar{P}}$, since moving the entire $\bar{\bar{P}}$ does not have any effect on the grammaticality of the forms involved. It has to be shown why in (8), as opposed to (9)-(11), we find the contracted forms of the preposition and the determiner. The constituent structures for these four example types are provided below.

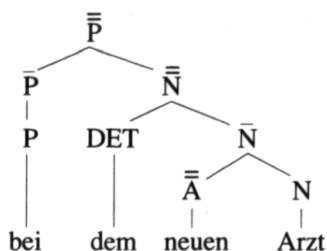
(17) (cf. 8)



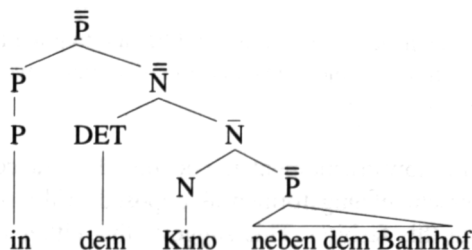
(18) (cf. 9)



(19) (cf. 10)

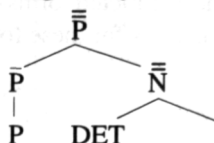


(20) (cf. 11)



Apparently, the crucial part of the structure (cf. 21) is exactly the same in all the diagrams listed above; therefore, it would not make any difference if one adopted one of the alternative definitions of c-command proposed within Government-Binding theory.

(21)



A further problem is related to the fact that contraction is obligatory under certain semantically definable circumstances. It is not really clear how Kaisse deals with this problem with regard to the English auxiliary 's in those cases where the full form is not allowed [8:42ff]. In all the explicit statements she makes concerning syntactic conditions on cliticization, she always uses the formulation "... *may* cliticize onto ..." [8:47, 53, 71, 72]. Finally, it should be mentioned that Hinrichs rejects the cliticization analysis for the Verschmelzungsformen on the basis of metatheoretical considerations related to the location of this rule in the overall grammar. Assuming an organization of grammar such as the one suggested by Zwicky and Pullum [17], all syntactic rules precede all cliticization rules, which implies that cliticization should not affect any syntactic rule. Since the rule of coordination is apparently sensitive to whether or not the article is attached, Hinrichs concludes that the contracted forms cannot be a consequence of cliticization

but are rather generated at an earlier stage in the grammar, cf. Hinrichs [7:134ff].

- (22) vor dem und nach dem Essen
 vor'm und nach'm Essen
 but: *vor dem und nach'm Essen

He concludes that the contracted forms "must be single constituents and not cliticized versions of prepositions." [17:135]

Summarizing this section, we can state that the *Verschmelzungsformen* cannot be interpreted as instances of cliticization, not only because they do not adhere to the general principles of simple cliticization as proposed by Kaisse [8], but also because they would violate the internal consistency of the grammar according to the theoretical framework in which they were proposed.

2.2. INFLECTION

The fact that contracted forms behave as single constituents prompts Hinrichs to conclude that they are to be analyzed as inflected forms. He points out that the *Verschmelzungsformen* generally behave more like inflected forms according to the criteria for distinguishing between inflection and cliticization provided by Zwicky and Pullum [17]. They show a "high degree of selection" since the articles are only attached to prepositions as is typical of affixes, but not of clitics; further, they show arbitrary gaps in the set of combinations, which is again allegedly typical for inflections. Zwicky and Pullum's claim that semantic idiosyncrasies are more characteristic of affixed words than of clitic groups seems consistent with the use of the contracted vs. noncontracted forms (cf. 1.2.).

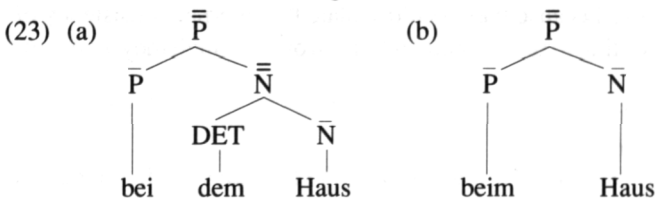
Once the decision is made that *Verschmelzungsformen* are to be analyzed as inflected forms, one faces two options: Either they are case-marked articles, or they are inflected prepositions. In their study of coalesced preposition-determiner forms in Italian, Napoli and Nevis [10] were confronted with a similar dilemma. After ruling out both a phonological and a cliticization analysis for forms like *nella* "in the," *sulle* "on the," *ai* "to the," etc., they decide to interpret them as inflected units in the lexicon. Both Hinrichs and Napoli and Nevis reject the option of analyzing such forms as case-marked articles: first, this would lead to a proliferation of cases; and second, it would lead to a number of problems with regard to syntactic structure. It would, for example, be a coincidence that *Haus* is a dative-marked noun in both *beim Haus* as well as *bei dem Haus* (cf. Hinrichs

[7:135]). Napoli and Nevis give a number of additional reasons why the "case-marked article analysis" should be rejected [10:11ff], most of which hinge on Government-Binding theory. The conclusion reached both by Hinrichs and Napoli and Nevis is to analyze the contracted forms as inflected prepositions.

The difficulties of analysis with regard to syntactic structure are solved by assuming two different phrase-structure rules to expand prepositional phrases:

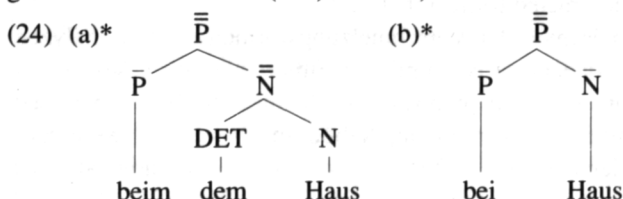
- (22) (a) $\bar{\bar{P}} \rightarrow \bar{P} \bar{\bar{N}}$
 (b) $\bar{\bar{P}} \rightarrow \bar{P} \bar{N}$

This results in the following structures for *bei dem Haus* vs. *beim Haus*:



However, the fact that $\bar{\bar{P}}$ takes an $\bar{\bar{N}}$ complement in (23a) but an \bar{N} complement in (23b) is somewhat problematical; Napoli and Nevis note that this segregates $\bar{\bar{P}}$ from all the other major categories [10:18]. Assuming the version of X-Bar theory that has been adopted within GB, the internal structure of phrases should pattern in similar ways. This, of course, raises the question of why $\bar{\bar{P}}$ should be different. According to Napoli and Nevis, this question cannot be satisfactorily answered.

The next problem that arises is to insure that the PS-rules in (22) will not generate structures like (24a) and (24b).



Napoli and Nevis make use of the feature $[\pm\text{inflected}]$ in their PS-rules. $\bar{\bar{N}}$ will then be introduced by $[-\text{inflected}]$, where as $\bar{\bar{N}}$ will be introduced by $[+\text{inflected}]$. Finally, Napoli and Nevis have to prevent $\bar{\bar{P}}$ s like **di la ragazza* "of the girl," instead of *della ragazza*. In Italian, the contracted forms are

obligatory if the preposition is monosyllabic except for the preposition *con* "with." They proceed to introduce a filter to exclude noncontracted forms [10:26]:⁵

a DS filter

*P ART

[-I]

where ART is the specifier of the NP object of the P

Exception: P=*con*

The specifier condition on the determiner is added in order to allow for the sequence *su il* in constructions such as:

Quando ho guardato in su, il medici mi ha spennellato la tonsile di iodio.

"When I looked up, the doctor painted my tonsils with iodine."

Here, of course, the contraction *sulle* is not possible. Contraction only occurs within prepositional phrases, which is true for German as well. If we adopt Napoli and Nevis' analysis to account for the German data, we could propose a filter such as (25).

(25) *P ART^{+def}
[-I]

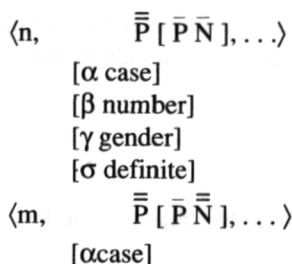
if used in a generic sense

where ART is the specifier of the NP object of the P

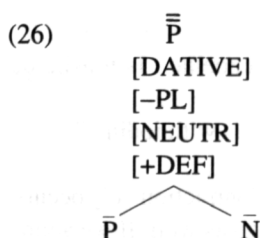
Exceptions: ART = *die, des, PL*

Before commenting on this analysis, I will first present Hinrichs' account of the German data. As already mentioned, he interprets Verschmelzungsformen as inflected prepositions. As a result of this analysis he needs to formulate two phrase-structure rules to expand \bar{P} . (cf. the discussion above). Since Hinrichs works within the GPSG-framework the formulation of parochial PS - rules is less problematic for him. Within this theory such strong predictions are not made with regard to the basic \bar{X} -scheme. Furthermore, Hinrichs does not have to introduce the diacritic feature $[\pm \text{inflected}]$. Instead, he assumes the following two PS-rules [7:136]:

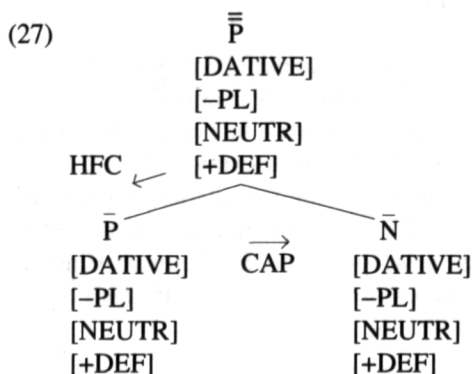
⁵The authors fail to mention that this filter applies only to monosyllabic prepositions.



The first rule will generate a structure such as (26):



Feature-values such as the ones listed in (26) are freely instantiated according to GPSG-theory; the information thus given will then “flow” within the structure according to certain principles that interact with each other to produce grammatical agreement, including the Head Feature Convention, which passes the features from the $\bar{\bar{P}}$ node to the \bar{P} node, and the Control Agreement Principle, which copies them from the \bar{P} node to the \bar{N} node.



The items *beim* and *Haus* can then be inserted under these nodes. On the other hand, generating a string like *bei dem Haus* does not require $\bar{\bar{P}}$ to have

grammatical information concerning gender, number, or definiteness. Case is the only syntactic property that crucially hinges on \bar{P} , and has to be passed to \bar{N} via \bar{P} .

Hinrichs does not mention how he would handle those problems that are taken care of by a filter such as (25) in a GB analysis. It seems, though, that he exceptions do not require any extra devices in the GPSG framework. All there is to say is that there simply are no lexical items available that could be inserted under nodes dominated by \bar{P} s such as the following:

(28) [+ GENITIVE]	[+ NOM]	[α CASE]
[+ MASC/NEUTR]	[+ FEM]	[β GENDER]
[+ SG]	[+ SG]	[- SG]
[+ DEF]	[+ DEF]	[+ DEF]

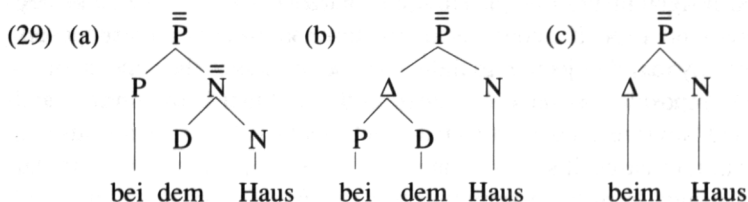
Finally, the restriction that the article has to be the specifier of the NP-object of the \bar{P} (cf. (25)) is automatically accounted for by the syntactic features encoded in the PS-rules.

Summarizing, we can state that the data is accounted for more straightforwardly in a GPSG framework, since no diacritic features such as [\pm inflected] or special filters need to be introduced. On the other hand, the principles employed in GPSG that "do the job," such as the Head Feature Convention or the Control Agreement Principle are somewhat *ad hoc* themselves, and they rely heavily on a specific constituent structure that has to be presupposed. However, my major objection to these analyses relates to the concept of an inflected preposition as such. Considering the fact that both Hinrichs and Napoli and Nevis reject the "inflected-article analysis" because it leads to a proliferation of cases, it seems strange that neither mentions that a similar problem results from the assumption that the grammar contains inflected prepositions. Recall that assuming that the Verschmelzungsformen are inflected articles leads not only to a proliferation of case, but also limits this proliferation to one single category: definite articles. Assuming inflected prepositions, on the other hand, leads to the supposition that every case-inflected category appears in all four cases except for the category preposition, the case of which will be lexically determined. Prepositions like *bei*, *zu*, *mit*, etc., are always marked for dative; prepositions such as *anstatt*, *während*, *wegen*, etc., are always marked for genitive; and others such as *ohne*, *für*, *gegen*, etc., are always marked for accusative. The defectiveness in all preposition paradigms severely violates the general concept of inflection. It is therefore concluded that the Verschmelzungsformen can be interpreted neither as instances of inflected articles, nor as inflected prepositions.

2.3. VERSCHMELZUNGSFORMEN AS COMPLEX CATEGORIES

2.3.1. Verschmelzungsformen as Primitive Lexical Entries

If we try to trace the basic shortcomings of Hinrichs' or Napoli and Nevis' account, we have to reconsider the question of why inflection was proposed in the first place. At one point, these authors find themselves compelled to decide on whether to analyze Verschmelzungsformen as case-marked articles or as inflected prepositions. The necessity of deciding between these two categories is clearly at the root of all the problems that follow from such a decision: proliferation of case, defectiveness of inflection paradigms, or general problems related to subcategorization. Encountering all these problems does not really come as a surprise since these forms are neither single articles nor single prepositions, but rather fused forms comprising both categories. This is exactly what the traditional term "Verschmelzungsformen" suggests. However, the overall organization of grammar these authors assume is not compatible with such a categorization: All lexical entries have to have a single category in order to be inserted under a certain node generated by PS-rules. Proposing the whole array of PS-rules given in (34) would carry the very function of this device *ad absurdum*:



And yet we find structures in natural languages that would license each one of these expansions, such as preposition-stranding in English for (a), and the contracted forms in German, Italian, French, and many other languages for (b) and (c). To claim that (29a) is the only legitimate PS-rule, and that structures such as (29b) or (c) are derived from this basic structure by grammatical operations, seems *ad hoc* and demands independent motivation. And there is, of course, the problem of how to label the categories dominating *bei dem* and *beim* (or *dem Haus*, for that matter). The label Δ in (c) cannot be a phrasal category since it immediately dominates a lexical node; on the other hand, it seems to correspond exactly to Δ in (b).

Since we identified PS-rules along with the notion of phrasal categories

as the cause of the problems in analyzing Verschmelzungsformen, I will now present a theory developed by Brame where both of these notions have been dispensed with. In his theory, words are not static entities, but dynamic functions operating on each other. A word is listed in the lexicon as an ordered n -tuple with the phonetic structure as initial coordinate, the intrinsic syntactic category as second coordinate, and the argument it operates on as third or higher coordinate. Before proceeding, we will give some examples:

- (30) $L_1 \langle \text{to}, P, D \rangle$
 $L_2 \langle \text{the}, D, N \rangle$
 $L_3 \langle \text{café}, N, 1 \rangle$

Thus, *to* is a preposition which selects a determiner; *the* is a determiner which selects a noun; and *café* is a functor which selects the identity category 1. Proceeding from there, Brame recursively defines the induced lexicon [2:142]:

- (i) If $L_1 \in \text{LEX}_0$, then $L_1 \in \text{LEX}$.
 If $L_1^n = \langle \chi, \phi, \psi_1, \dots, \psi_n \rangle \in \text{LEX}$ and $L_j^m = \langle \gamma, \psi_1 \sigma, \phi_1, \dots, \phi_m \rangle \in \text{LEX}$,
 for $n \geq 1, m \geq 0$, then
 $\langle x - y, \phi \psi_1 \sigma, \phi_1, \dots, \phi_m, \psi_2, \dots, \psi_n \rangle \in \text{LEX}$.

Thus, two words can be linked if the argument category of one of them is identical to the intrinsic category of the other. Hence, we can apply *to* to *the*, which gives us $\langle \text{to-the}, PD, N \rangle$. *To-the* is, thus, a complex category which, in turn, can be applied to *café* rendering $\langle \text{to-the-café}, PDN, 1 \rangle$ as its value. Clearly, we can apply *the* to *café* first, which gives us $\langle \text{the-café}, DN, 1 \rangle$ and then apply *to* to this complex category from the left, since the argument structure of *to* is only required to link to the initial subcategory of the intrinsic category of $\langle \text{the-café}, DN, 1 \rangle$ (see part (ii) of the definition above). The associativity of the system rejects the premise that a phrase structure such as (29a) is in any way more legitimate or more basic than the structure given in (29b).

On the basis of the recursive definition given in (31), together with appropriately specified base words, infinitely many well-formed composite words can be induced; any ill-formed strings will be due to carelessness in stating the domain to which a given lexical entry applies. For example, the string *to-café* is ungrammatical since prepositions do not operate on nouns; the string *to-the*, on the other hand, is well-formed, a claim which is corroborated by the fact that a corresponding sequence is lexicalized in many languages as, for example, French *au* (*à le*), Italian *sul* (*su il*), or the Verschmelzungsformen

in German. Similarly, we expect to find lexical entries that correspond to the well-formed string *to-the-café* and, in fact, there are primitive items like *home*, which have the same intrinsic category, namely PDN (cf. [1:323]). The difference between items like *au* or *home* as opposed to *à-la*, or *to-the-café* is that the former are primitive lexical entries, whereas the latter are lexically derived. This suggests a constraint on primitive lexical items that Brame formulates as follows:

“A primitive lexical entry must correspond to some composed set of lexical categories.” [1:323]

We will thus analyze *beim* as a primitive lexical entry with a complex category structure, and *bei-dem* as an induced form. The specification of the intrinsic syntactic category of *beim* vs. *bei-dem* should differ in order to reflect the difference in distribution noted in § 1. This leads to the question of whether the argument structures of both items reflect this distinction as well. We will come back to this problem when we discuss the details of the categorial structure in (2.3.2).

Summarizing, we can state that the account given here does not rely on any devices such as PS-rules, nor on transformations (cliticization), nor on principles like the Head Feature Convention or the Control Agreement Principle. No filters are employed to rule out ungrammatical forms such as **bein*, **währends*, etc., since these forms simply are not contained in the finite list of primitive lexical entries. Verschmelzungsformen are not looked upon as marked phenomena (cf. [16:289]), causing all kinds of disturbances and irregularities in the overall organization of the grammar; quite the contrary, forms like these are exactly what we expect to find in natural languages as long as their categorial structure corresponds to an induced set of lexical categories in the respective language.

2.3.2. *Verschmelzungsformen as Induced Lexical Items*

Before finishing the discussion of Verschmelzungsformen in German, I will try to develop one further analysis within the framework sketched in the preceding section. There are two problems with the analysis given in (2.3.1). First, although the ‘gaps’ discussed in § 1.1.2 can be accounted for by their non-inclusion in the primitive lexicon, they do not seem to be totally random, and the pattern in which preposition-article pairs fuse together might be accounted for by deriving the Verschmelzungsformen from primitive components. Second, the paradigm given in (31) suggests that the fused forms in German do have some internal structure as opposed to, say, French /o:/ *au*.

- | | | | |
|------|----------------|------------------|-----------------|
| (31) | zu gutem Wein | zu guter Butter | für gutes Geld |
| | zu diesem Wein | zu dieser Butter | für dieses Geld |
| | zu jedem Wein | zu jeder Butter | für jedes Geld |
| | zu dem Wein | zu der Butter | für das Geld |
| | zum Wein | zur Butter | fürs Geld |

Focusing only on the determiners, I will suggest that they be analyzed as *de-m*, *de-r*, and *da-s*, respectively. For the determiners proper, I will assume lexical entries such as:

- (32) $\langle \text{de}, D, N \rangle$
 $\langle \text{da}, D^{\text{da}}, N \rangle$

Hence, determiners do not carry any information with regard to gender, number, or case. I take *m*, *r* and *s* to be members of the suffix lexicon which apply to words according to the rule of suffixation [2:155]:

- (33) If $L_j = \langle \chi\gamma, \phi\sigma, \psi_1, \dots, \psi_n \rangle \in \text{LEX}_0$ and $L^s = \langle \gamma, \phi, \sigma \rangle \in \text{LEX}^{\text{suf}}$,
 then $\langle \chi\gamma, \phi\sigma, \psi_1, \dots, \psi_n \rangle \in \text{LEX}_0$.

I regard the suffixes mentioned above as complex units operating on determiners:

- (34) $\langle m, \text{masc sg dat}, D \rangle$
 $\langle r, \text{fem sg dat}, D \rangle$
 $\langle s, \text{neut sg acc}, D^{\text{da}} \rangle$

We can then induce determiners like *de-m*, *de-r* and *da-s*. At this point, the question arises as to just what prevents us from applying *D* to *N*, yielding ungrammatical strings like **de-Wein*, **da-Geld*, etc. While it is true that *D* selects *N*, there aren't any *N*s available in the lexicon in German; there are only *N*^{gend}s. In other words, every noun contains an inherent gender-category. Thus, we get the following lexical entries:

- (35) $\langle \text{Wein}, N^{\text{masc}}, 1 \rangle$
 $\langle \text{Butter}, N^{\text{fem}}, 1 \rangle$
 $\langle \text{Geld}, N^{\text{neut}}, 1 \rangle$

Before *D* can select *N*, the morphological marker must have applied to it. The application of the morphological marker supplies not only gender, but number and case as well. In order to match the morphological structure of

determiners, nouns have to be operated on by number and case markers. Thus, a form like *Kindern* is induced in two steps:

$$(36) \langle \text{Kind}, N^{\text{neut}}, 1 \rangle \circ \langle \text{er}, \text{pl}, N^{\text{neut}} \rangle = d\langle \text{Kind-er}, N^{\text{neut pl}}, 1 \rangle \\ \langle \text{Kind-er}, N^{\text{neut pl}}, 1 \rangle \circ \langle \text{n}, \text{dat}, N^{\text{gend pl}} \rangle = \langle \text{Kind-er-n}, N^{\text{neut pl dat}}, 1 \rangle$$

We suggest that the morphological markers applying to determiners are primitive units with a complex category, whereas the inflectional morphological structure of nouns is the result of lexical composition. Furthermore, the categorial structure of *Kind-er-n* licenses the specific sequence given for the categories of the suffixes in (34) (cf. Brame's constraint on primitive lexical entries).

Summarizing the exposition to this point, we assume the following structure of "noun-phrases" in German:

$$(37) \begin{array}{c} \overbrace{\text{de} - \text{m}} \\ \downarrow \quad \downarrow \\ \text{de} - \text{m} - \text{Wein} \end{array}$$

Now the question arises as to how prepositions link up to this structure. The traditional classification of prepositions in German as accusative-, dative- and genitive-prepositions suggests already that prepositions select certain cases. Thus, we assume the following lexical entries:

$$(38) \begin{array}{l} \langle \text{für}, \text{P}, \text{acc} \rangle \\ \langle \text{mit}, \text{P}, \text{dat} \rangle \\ \langle \text{wegen}, \text{P}, \text{gen} \rangle \end{array}$$

A preposition-determiner syntagma has, accordingly, a structure such as the following:

$$(39) \begin{array}{c} \overbrace{\text{zu} - \text{de}} \\ \downarrow \quad \downarrow \\ \text{zu} - \text{de} - \text{m} \end{array}$$

Following Brame in his assumption that suffixes operate on words, but not vice versa [2:155], *zu* can only have operated on the morphological marker *m* after the string *de-m* has been induced. Therefore, we must adjust the argument structure of prepositions accordingly:

$$(40) \langle \text{zu}, \text{P}, \text{D}, \text{gen}, \text{dat}, \text{num} \rangle$$

In the case of *Verschmelzungsformen* we would assume a phonetically empty determiner which is operated on by a morphological marker before being selected by a preposition. Skepticism is certainly warranted when

phonetically empty elements are introduced. But even though this determiner does not have any phonetic content, it has a clear syntactic function. Stating the lexical entry of this determiner explicitly is, however, problematic. The first question to consider is whether the feature [\pm generic] is to be included in its intrinsic category, a possibility which might have to be ruled out *a priori*:

“Generic interpretations of sentences result in the most common cases from the interaction between the tense and aspect, and the interpretation of the subject noun phrase (NP). Hence, the generic interpretation of an expression is not determined, at any lexical level, in isolation from its context.” [13:123]

However, in the dialect of Amern mentioned above, we find two distinct sets of determiners which bring about a lexical disambiguation between a generic and a nongeneric reading [7:94]:

- (41) æt pɛʁt löpt flɔtər əls dər honk
 ‘Das Pferd läuft schneller als der Hund’ (generic reading)
 dɔt pɛʁt löpt flɔtər əls də honk
 ‘Das Pferd läuft schneller als der Hund’ (specific reading)

In a similar way, the following two sentences are disambiguated in Standard German:

- (42) Ich gehe gerne ins Kino. (generic reading)
 Ich gehe gerne in das Kino. (Specific reading)

Here, too, the distinction between a generic and a specific reading manifests itself on a lexical level. As has been pointed out above, the distribution of the two sets of determiners in the dialect of Amern exactly parallels the distribution of the contracted versus noncontracted forms in Standard German. This parallelism should be reflected in the respective lexical entries of the determiners. Hence, we have to account for the fact that the generic determiner is incompatible with any kind of modification, be it restrictive relative clauses, attributive adjectives, or prepositional phrases. Brame has argued that restrictive relative clauses are lexically linked to determiners [2:166]. If this is true, one might further investigate the possibility of any kind of restriction being lexically linked to the determiner in a (ideally) unique way. For the time being, we will assume the existence of a certain ‘specifier-function’ F^{spec} which, in turn, selects the respective modifier. The

primitive specification of determiners can be tentatively given as follows:

- (43) $\langle \text{de}, D^{\text{spec}}, N, F^{\text{spec}} \rangle$
 $\langle \text{da}, D^{\text{spec}}, N, F^{\text{spec}} \rangle$
 $\langle \Lambda, D^{\text{gen}}, N \rangle$

Accordingly, in the local dialect of Amern we get lexical entries such as:

- (44) $\langle \text{d}\text{ə}, D^{\text{spec}}, N, F^{\text{spec}} \rangle$
 $\langle \text{ə}, D^{\text{gen}}, N \rangle$

(We assume here that the neuter determiners are composed lexical functions $\text{d}\text{ə}-t$, $\text{ə}-t$, where the t carries the morphological information.)

At this point, there is still one outstanding problem: How do we make sure that the phonetically empty determiner is always operated on by prepositions in order to prevent the induction of strings like *m-Wein*? This problem is admittedly rather intricate, and it cannot be solved here in a satisfactory way. I would nonetheless like to point out some advantages that would result from analyzing Verschmelzungsformen as lexically induced forms. It appears that this analysis allows us to account for the 'gaps' in a rather straightforward manner. That is, we cannot induce forms like **fürie* since the determiner *die*, as opposed to *de-r da-s*, etc., is a primitive lexical entry and there is no suffix /i:/ available in the lexicon. Our reasoning is supported by the fact that in dialects that have /də/ instead of /di:/, we find Verschmelzungsformen such as /Inə/ 'in die', /awfə/ 'auf die', etc. In these dialects the determiner /də/ is not a primitive lexical entry, but is composed on the basis of paradigmata such as (45):

- (45) für gutə Butter
 für də Butter

The nonoccurrence of forms like **wegens*, **währends*, etc., can be explained in a similar fashion; the paradigm in (46) suggests that the genitive determiner *des* is a primitive lexical entry:

- | | |
|------------------|---------------------------|
| (46) zu dem Wein | wegen des Wetters |
| zu gutem Wein | wegen schlechten Wetters |
| | *wegen schlechtes Wetters |

We find further evidence in favor of the nonderived status of *des* if we

compare its phonological structure to that of induced determiners:

- (47) [de:m]
 [de:ɐ] [dɛs]
 [das]

Whereas [de:] and [da] are phonologically well-formed word sequences (cf. *Tee*, *da*, etc.), [dɛ] is not, since [ɛ] is not an admissible word-final vowel. Analyzing *des* as a primitive lexical entry explains why we cannot induce forms like **wegens*: again, ⟨s, masc sg gen, D⟩ is not available in the lexicon.

The nonoccurrence of Plural Verschmelzungsformen is more difficult to explain. It might be related to the fact that there is no definite plural determiner which allows a generic interpretation:

- (48) Ein Tiger hat Streifen.
 Der Tiger hat Streifen.
 Tiger haben Streifen.
 *Die Tiger haben Streifen.

The main point is that we wish to relate directly the fact that only the determiners *der*, *das*, *dem* and *den*, as opposed to *des* and *die*, have 'fused' counterparts, to the observation that the former appear to be lexically induced whereas the latter are primitive lexical entries. This is accomplished by the analysis provided here. However, as has been pointed out above, it is not clear whether this analysis can be formalized within the theory of recursive categorial syntax. Hence, it might turn out that the problem related to the empty determiner addressed above cannot be solved; in other words, that we won't be able to state the lexical entries in such a way that any operation in the lexicon would result in a grammatical string. Since we take the idea seriously that shortcomings in formalizing a grammar reflect substantial problems of the theory as such, we might have to conclude that the analysis of Verschmelzungsformen as productively composed lexical entries is to be abandoned in favor of the analysis given in (2.3.1).

REFERENCES

1. Brame, M. K. 1981. The head-selector theory of lexical specification and the nonexistence of coarse categories. *Linguistic Analysis* 10(4).
2. Brame, M. K. 1985. Recursive categorial syntax II: n-arity and variable continuation. *Linguistic Analysis* 15:137-176.
3. Carlson, G. 1977. Reference to kinds in English. University of Massachu-

- setts at Amherst doctoral dissertation. Distributed by Indiana University Linguistics Club.
4. *Duden 9. Zweifelsfälle der deutschen Sprache* 1965. Bibliographisches Institut Mannheim/Wien/Zürich: Dudenverlag.
 5. Gazdar, G. and G. Pullum. 1982. *Generalized phrase structure grammar: a theoretical synopsis*. Indiana University Linguistics Club.
 6. Heinrichs, H. M. 1954. *Studien zum bestimmten Artikel in den germanischen Sprachen*. Gießen: Wilhelm Schmitz Verlag.
 7. Hinrichs, E. 1984. Attachment of articles and prepositions in German: simple cliticization or inflected prepositions. *Ohio State University Working Papers in Linguistics* 29:127-138.
 8. Kaisse, E. 1985. Connected speech: the interaction of syntax and phonology. New York: Academic Press.
 9. Matthias, T. 1914. *Sprachleben und Sprachschäden*. Leipzig.
 10. Napoli, D. J. and J. Nevis. (Forthcoming). Inflected prepositions in Italian.
 11. Sanders, D. 1897. *Wörterbuch der Hauptschwierigkeiten in der deutschen Sprache*. Berlin: Langenscheidtsche Verlags-Buchhandlung.
 12. Schaub, S. 1979. Verschmelzungsformen von Präpositionen und Formen des bestimmten Artikels. *Phonologische Probleme des Deutschen*, ed. by Heinz Vater, 63-96. Narr: Tübingen.
 13. Ter Meulen, A. 1986. Generic information, conditional contexts and constraints. In: *On conditionals*, ed. by E. Closs Traugott et al. Cambridge: Cambridge University Press.
 14. Zwicky, A. 1977. On clitics. Indiana University Linguistics Club.
 15. Zwicky, A. 1982. An expanded view of morphology in the syntax-morphology interface. *Reprints of the plenary session papers*. XIIIth International Congress of Linguists, Tokyo.
 16. Zwicky, A. 1985. Clitics and particles. *Language* 61:283-305.
 17. Zwicky, A. and G. Pullum. 1983. Cliticization vs. inflection: English *n't*. *Language* 59:502-513.